AL/CF-TR-1994-0052

ARMSTRONG LABORATORY

THE EFFECT OF INTEROCULAR DISTANCE UPON DEPTH PERCEPTION WHEN USING STEREOSCOPIC DISPLAYS TO PERFORM WORK WITHIN VIRTUAL AND TELEPRESENT ENVIRONMENTS

Louis B. Rosenberg

CENTER FOR DESIGN RESEARCH
STANFORD UNIVERSITY
STANFORD CA 94305

JULY 1992

INTERIM REPORT FOR PERIOD JANUARY 1992 TO JUNE 1992

AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-7901

TECHNICAL REVIEW AND APPROVAL

AL/CF-TR-1994-0052

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Regulation 169-3.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Thomas J. Moore

THOMAS J. MOORE, Chief
Biodynamics and Biocommunications Division
Crew Systems Directorate
Armstrong Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
July 1992

3. REPORT TYPE AND DATES COVERED
Interim for January 1992 to June 1992

4. TITLE AND SUBTITLE
The Effect of Interocular Distance upon Depth Perception when Using Stereoscopic Displays to Perform Work within Virtual and Telepresent Environments

5. FUNDING NUMBERS
PE - 61101F
PR - 7231
TA - 38
WU - 08

6. AUTHOR(S)
Louis B. Rosenberg

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Design Research
Stanford University
Stanford CA 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory, Crew Systems Directorate
Biodynamics and Biocommunications Division
Human Systems Center
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7901

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/CF-TR-1994-0052

11. SUPPLEMENTARY NOTES

Support for this work was provided in part through the Air Force Office of Scientific Research (AFOSR) Laboratory Graduate Fellowship Program (LGFP).

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Design parameters must be chosen for stereoscopic vision systems intended for telepresence or virtual environments. It is not clear that parameters that closely match human physiology optimize operator performance or enhance understanding of displayed information. This study investigates interocular distance, a basic parameter governing stereopsis, and assesses its effect upon operator performance in a binocular depth perception task.

A peg alignment task was used to quantify operator performance. Subjects were required to align virtual pegs presented by a shuttering stereoscopic display. Interocular distance of the projection model was varied between 0 cm and 8 cm. Results revealed a tenfold performance improvement when stereoscopic projections replaced pure monocular projections. A roughly logarithmic relation was found between interocular distance and operator performance. As interocular distance increased from 0 cm to 2 cm, performance improved rapidly then asymptotically approached a maximum value. Although average physiological interocular distance is 6.3 cm, no measurable increase in performance was found for interocular distances greater than 3 cm.

There is no measurable advantage to using anthropomorphic interocular distances when projecting stereoscopic images. Because reduced parallax in these projections typically increases displayable depth range, reduces image fusion problems, and reduces operator fatigue, there is a strong motivation to use the smallest possible interocular distance in the projection of stereoscopic images. This study suggests that the use of interocular distances between 2 cm and 3 cm will not degrade performance in depth perception tasks and may facilitate the projection of stereoscopic images.

14. SUBJECT TERMS
stereoscopic displays
virtual reality
stereopsis
stereo vision
depth perception
interocular distance
telepresence

15. NUMBER OF PAGES
30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Background

It is known that binocular stereopsis enhances performance in depth perception tasks [1]. Although much depth information can be inferred from monocular depth cues alone, stereoscopic depth cues provide additional information that often enhances the speed and accuracy of tasks requiring depth perception. When considering the use of stereoscopic vision as part of the user interface for telepresence or virtual environment systems, questions arise regarding the design parameters for implementing artificial stereopsis. Visual projection technologies used in telepresence and virtual environment systems offer new freedom from the biological constraints on stereoscopic vision. Many of the parameters of human vision which could not have been optimized or even altered in the past have suddenly become design parameters. This study investigates the most basic parameter of stereoscopic vision, interocular distance, and assesses its effect upon performance in basic depth perception tasks. Although average physiological eye separation is 6.3cm [2], it is unclear whether the use of such a typical value yields maximal performance in depth perception tasks. The purpose of this study is to provide answers to questions such as "How much stereopsis is enough?" and "How much stereopsis is too much?" by developing relations between interocular distance and performance. Once we get a firm grasp on the effect that interocular distance has upon operator performance, we can develop guidelines for maximizing the performance of operators using stereoscopic vision systems for telepresence and virtual environment systems. A sound understanding of the relationship between interocular distance and performance could even help fine tune a perceptual environment to enhance operator performance for a particular task.

Previous studies have presented conflicting results over the advantage of stereoscopic vs monocular projections used for telepresence systems. Many studies have shown that stereoscopic displays, as compared to monocular displays, do not provide significant performance advantage [3, 4, 5, 6]. Other studies have indicated that performance associated with stereoscopic displays

was greatly superior to monocular displays under most conditions tested [7, 8, 9, 10, 11]. This study attempts to gain deeper insight into the usefulness of stereo projections by comparing monocular and stereo projections *not* as binary alternative conditions but rather by comparing a full range of interocular distances from pure monocular to exaggerated stereo.

Stereoscopic depth perception is primarily the result of differences in the perspective viewpoints incident on each eye. Because differences in left and right vantage points are entirely dependent upon interocular distance, interocular distance is the primary parameter governing stereopsis. The greater the distance between the eyes, the greater the difference in the perspective incident upon each eye, and thus the stronger the stereoscopic effect. As interocular distance goes to zero, all differences in perspective viewpoint are lost, and all stereoscopic depth cues disappear. By varying the interocular distance in a simple depth perception performance test, we can develop a relation between the degree of stereopsis and operator performance for the range of vision from pure monocular to exaggerated stereo. Although it might seem strange to vary a parameter which is usually fixed by human physiology, interocular distance can be varied freely when generating artificial stereoscopic images. Before describing the details of the experiment, it would be best to review the basic theory behind the creation and projection of stereoscopic images and clarify how variation of interocular distance fits into the projection model.

Stereoscopic Images

The perception of all stereo images, whether real or artificial, follows the same basic process in the visual system. When a viewer looks at a real object at some distance in the visual field, the left eye and the right eye are presented with slightly different perspective viewpoints. As a result, the images projected on the retina of each eye will not be identical. The primary difference between the image projected on each retina is a small horizontal offset known as *lateral retinal image disparity.* Lateral retinal disparity is defined as the difference in relative position of the visual images of an object on the two retinas due to the lateral separation of the eyes [1]. The visual system in the cerebral cortex has receptive fields sensitive to lateral retinal

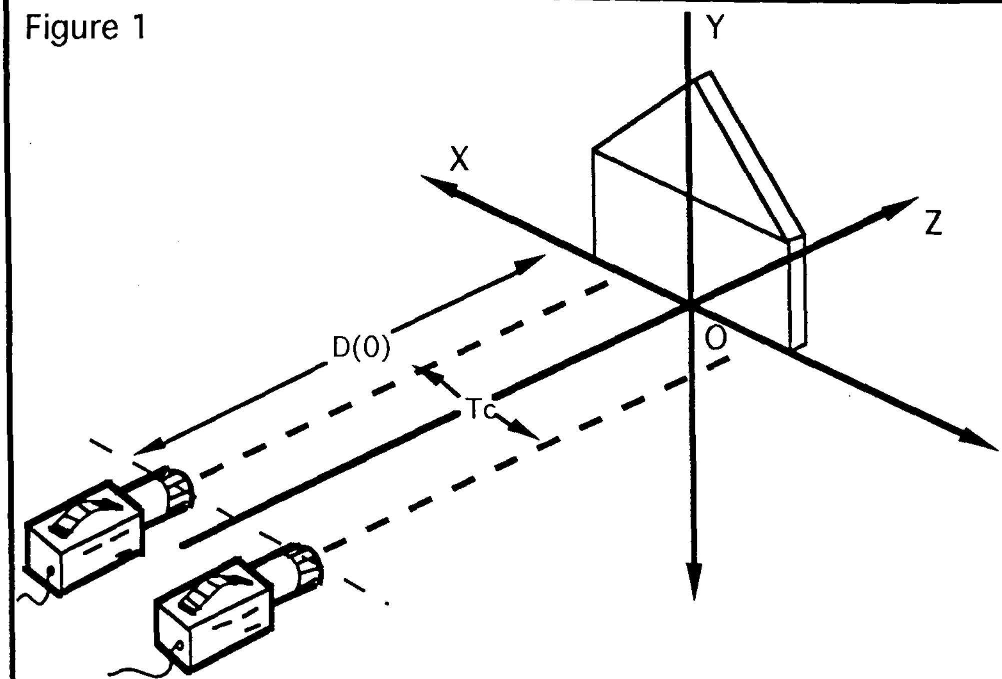
disparity and codes this information into a sense of depth. The brain's ability to merge the information gathered by each eye into a single meaningful image that contains depth information is called image *fusion*. The human vision system does not transduce depth absolutely like a range-finder, but rather compares the retinal disparities of the various objects in the visual field to get relative depth between objects. In addition to stereo cues, many monocular cues such as relative size, shading, motion parallax, perspective, and interposition are used by the brain to infer depth information.

Stereo images with accurate binocular depth cues can be produced by presenting each of the operator's eyes with slightly different views of an object in the same way the operator would perceive the real object. Stereo vision systems are commercially available which provide a means of projecting images independently to each eye. To generate an accurate stereoscopic image, a mathematical model is required to generate the appropriate left eye and right eye projections for a given vantage point. The following section discusses the details of the particular projection model used in this study.

The Stereoscopic Model

If we think of the geometry of the human visual system as two parallel video cameras spaced a distance **Tc** apart, a simple mathematical model for projecting stereoscopic images can be developed. By placing an object at distance **D(0)** in front of the camera pair as shown in Figure 1, each camera will see a slightly different perspective viewpoint of the object.

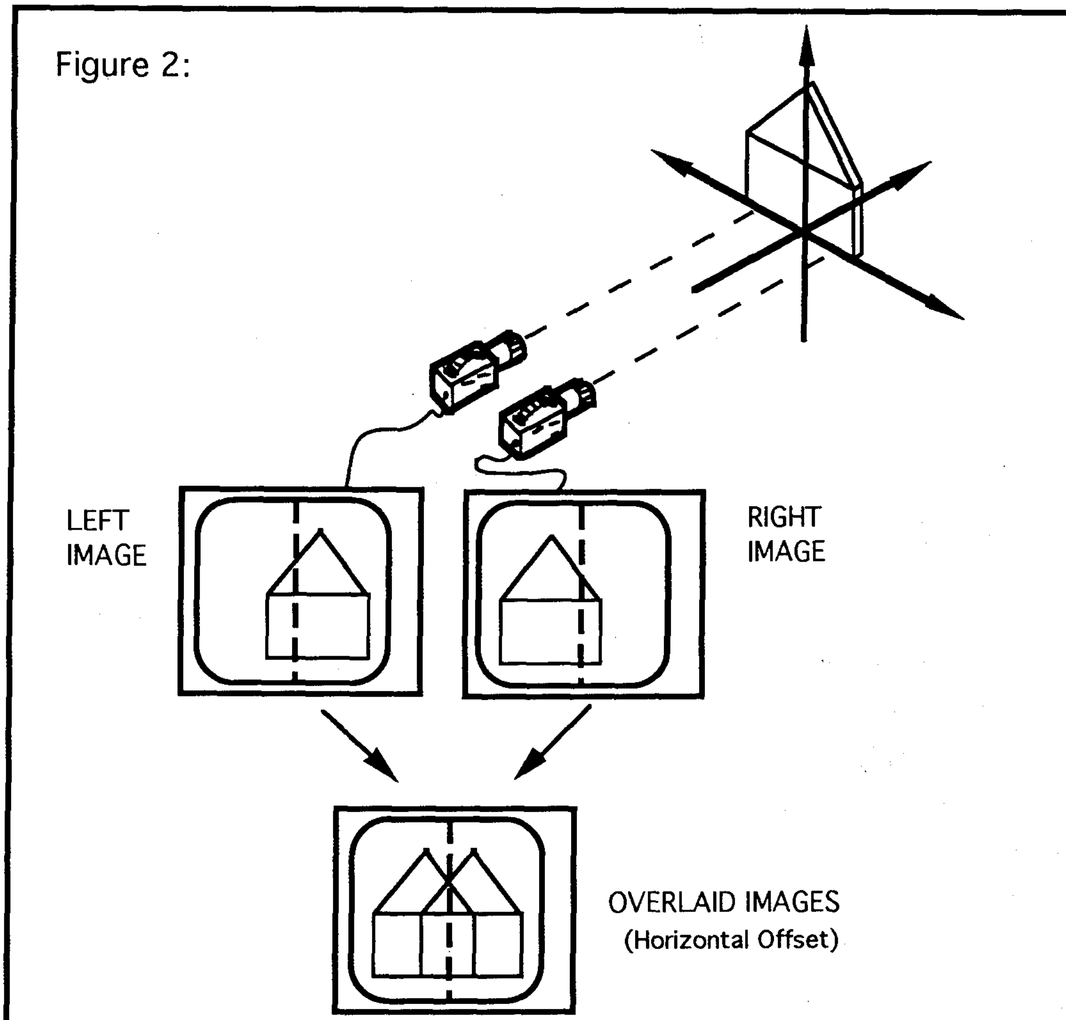
Figure 1



Dual camera model assuming cameras are fixed along Z axis pointing in the positive Z direction. D(0) is the distance from cameras to the XY plane. Tc is the distance between the parallel lens axes. Object is located on XY plane.

Figure 1: Dual Camera Model of Stereopsis

If we overlaid the video signal from each camera on the same video monitor, the left and right images would appear very similar but would be offset from each other by a small horizontal distance as shown in Figure 2. The offset distance depends entirely on the ratio of **Tc** to **D(0)**. If **Tc** is held constant and a number of objects are presented in the visual field at different depths **D(i)**, each object would produce a different horizontal offset which corresponds to the depth of that object. If we think of the pair of planar images as a means of storing the depth information, the horizontal offsets between the left and right images of each object are the primary method of coding the depth of that object.



Since each camera has a different perspective view point of the object, left and right images are offset horizontally. If we overlay left and right images we can clearly see horizontal offset analogous to retinal disparity in the human visual system.

Figure 2: *Dual Camera Model* : Horizontal Offset Between Left and Right Images

Rather than overlay the left and right images on a single monitor and measure the horizontal offsets to yield depth information as described above, we can project each image separately to the user and depend on the human visual system to decode the scene. When presented with a binocular image pair, the human brain will try to fuse the two flat images into a single stereoscopic perception rich in depth information.

Video vs Graphics for Stereo Projection

The projection of the images gathered by a stereo camera pair directly to the eyes of a user, as described above, is often used in telepresence vision systems to convey stereoscopic information. Although this dual camera technique is an important part of many telepresence systems, it was *not* used for this study because the physical hardware would have greatly limited parameter variation. In order to vary interocular distance (**Tc**) in a dual camera system, the physical separation between the two cameras would have to be altered. Such an alteration would have made the rapid testing of random interocular distance trials impossible.

Rather than projecting real video images from real cameras, the same effect can be produced using a high fidelity graphics computer to generate simulated images for each eye. To produce a computer generated stereoscopic image, we simply need to produce graphical binocular images similar to those produced by parallel video cameras (i.e., objects of a particular depth have a particular horizontal offset between the left and right images). Before describing the particular method used to project graphical stereoscopic images, a quantity known as *parallax*, representing the horizontal offset between left and right images, needs to be introduced.

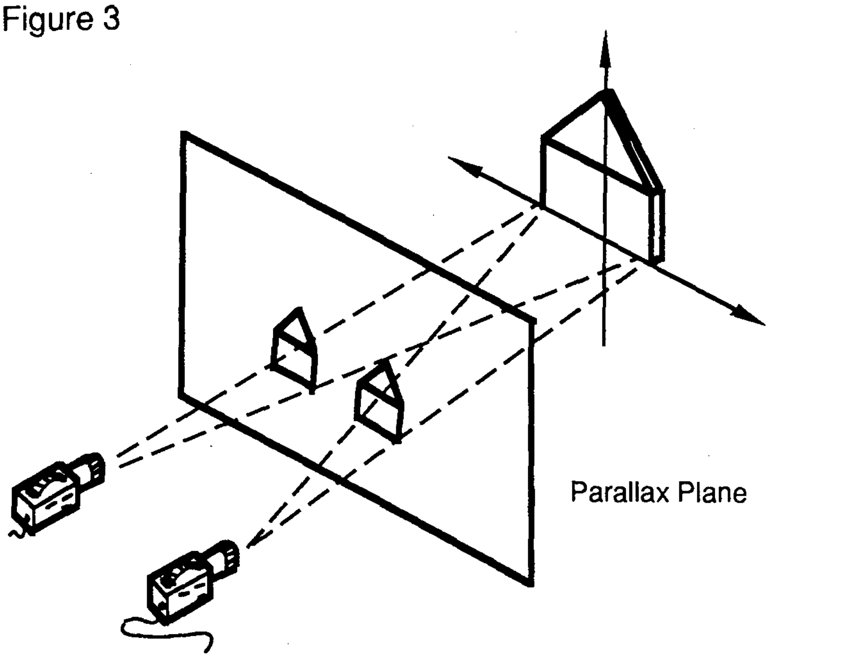
The Concept of Parallax

Although biological stereopsis is usually discussed in terms of lateral retinal disparity, when discussing artificial stereo projections it is convenient to introduce a quantity called parallax. Parallax, like disparity, is a horizontal offset between left and right images. The difference is that while disparity is measured at the retina, parallax is measured at some arbitrary plane between the eyes and the object [12].

Parallax is easily understood by imagining that you are looking at an object through a window. Assume for now that the window lies at some distance between your eyes and the object as depicted in Figure 3. If you could close one eye and trace the image as you see it passing through the plane of glass, then close the other eye and trace the new image you see it passing

through the glass, you would get the outlines of two images which were offset horizontally. This offset is parallax. Parallax, just like lateral retinal disparity, is dependent upon the ratio of interocular distance and distance to the object. Unlike lateral retinal disparity, parallax is also dependent on the location of the chosen parallax plane (i.e., the location of the window). For example, if you moved the plane of glass closer to your eyes and traced the same object as before, the horizontal offset between the left and right images would increase. If you moved the plane of glass closer to the object, the horizontal offset would go to zero.

Figure 3



Imaginary plane of glass between object and vantage point demonstrates the concept of the parallax plane. Diverging dotted lines demonstrate that offset depends on location of parallax plane.

Figure 3: The Concept of Parallax

Any object whose depth corresponds to the depth of the chosen parallax plane has zero horizontal offset between the left and right images. Thus we define the chosen parallax plane as the *plane of zero parallax*. If we consider a visual field with numerous objects at different depths and pick an arbitrary but fixed parallax plane, some objects will fall in front of the plane, some will

fall behind the plane and some will fall on the plane. The greater the distance an object is behind the plane of zero parallax, the greater the *positive parallax*. The greater the distance an object is in front of the plane of zero parallax, the greater the *negative parallax*. Negative parallax is often called crossed parallax because the left and right images flip sides.

Why introduce this arbitrary reference plane and defined parallax values relative to this plane? The answer has to do with the means of image projection. Most methods used to present the left and right images to the eyes do not project the image directly on the retina, but rather project the image on a screen which is some distance away from the retina. Thus, rather than deal with lateral retinal disparity directly, it is more convenient to deal with parallax at the plane of the video monitor.

Projection Hardware

The particular method used to generate stereoscopic images in this study presented stereo images on a single video monitor located 80 ± 4 cm from the user. In order to present different images to the left and right eyes using a *single* video monitor, shuttering stereoscopic glasses were used. CrystalEyes liquid crystal shuttering glasses allow the rapid alternation of two images on a single monitor while ensuring that each alternating image reaches only the intended eye. The shutters, synchronized to the monitor's raster scan, rapidly block and unblock alternate eyes when the appropriate image is displayed [13]. The left and right images are flashed at 120 Hz, which is fast enough that no flicker is noticeable to the user.

If we consider the screen of the monitor as the plane of zero parallax, we can generate stereoscopic image pairs with zero parallax, positive parallax, or negative parallax. If we generate left and right images on the screen which have no parallax, the image pair will have no horizontal offset, and the image appears to be located at the depth of the screen surface. If we produce images on the screen with positive parallax, the images will appear to be behind the screen surface. If we produce images on the screen with negative parallax the images will appear to be in front of the screen surface. Thus to

place images anywhere on the z axis, we simply define the horizontal offset between the left and right images when projected at the plane of the screen.

Stereo Perspective Projections

Whereas a *monocular perspective projection* produces a single rendering of a three dimensional object on a flat screen, a *stereo perspective projection* produces a left-right pair of renderings that represent the object at some depth in front of or behind the plane of the screen. A monocular perspective projection is achieved by considering a single vantage point, known as the center of the projection. The projection method is best understood by imagining the object to be at its desired location in three dimensional space and by pretending to sweep a line from the center of the projection to every point on the object. The planar projection of the object is achieved by locating the intersections of the sweeping line with the object and plotting those points at the locations where the sweeping line passes through the plane of the screen. The result is a planar description of the three dimensional object as would be perceived by a single eye at the center of projection. A stereo perspective projection is achieved using the same method but by choosing a different vantage point for the left and right projections such that centers of projection for the left and right images are separated by the desired interocular distance. Thus stereoscopic images modeled with arbitrary interocular distance can be generated by varying the distance between the left and right centers of projection [14].

EXPERIMENTAL PROCEDURE

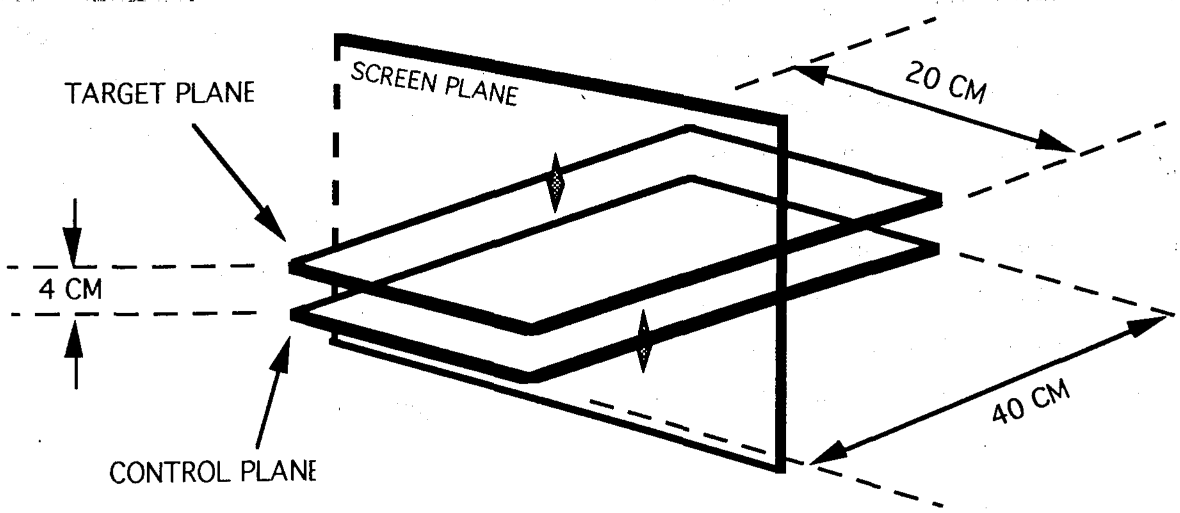
To investigate the effect that interocular distance has upon user performance in simple depth perception tasks, the following experiments were developed. Subjects were required to visually align small pegs in three dimensional space. Performance in these peg alignment tasks was recorded as error in peg alignment. All tests used the CrystalEyes liquid crystal shuttering glasses in conjunction with a Silicon Graphics graphical display to present virtual pegs to the subjects. The use of graphical simulation for these peg

alignment tests allowed for rapid variation of interocular distance between trials without the subjects being aware that any change had been made.

Test Set-Up

Each subject was outfitted with liquid crystal shuttering glasses and seated 80 cm from the face of a single stereo display monitor. Subjects were presented with a simple stereoscopic image that consisted of two small pegs on a solid blue background. Both pegs were modeled identically as diamond shaped polygons 3.5 cm high and 0.8 cm wide. The pegs were rendered and shaded three dimensionally as realistic solid objects to provide monocular depth cues in addition to the stereo cues. The use of simple, perceptually rich figures provided a controlled but realistic perceptual environment for testing. These monocular cues include linear perspective, perspective change, and relative size change. One of the pegs was defined as the *target peg* and was placed by the computer at a random location in a plane called the TARGET X-Z plane (a horizontal plane into the monitor). The other peg was defined as the *control peg*. and was positioned by the subject using a standard mouse interface. The subject could move the control peg anywhere in a plane parallel to the TARGET X-Z plane called the CONTROL X-Z plane (a horizontal plane into the monitor). These two parallel planes were defined identical in size, being 20 cm wide and 40 cm deep as shown in Figure 4. The TARGET X-Z plane was positioned 2 cm above the center point of the screen and the CONTROL X-Z plane was positioned 2 cm below the center point of the screen. Restricting peg motion to these parallel planes guaranteed that the bottom of the target peg would always be 0.5 cm above the top of the control peg and thus eliminated vertical displacement between the pegs as a variable in this study.

Figure 4



Target Plane and Control Plane shown in relation to the plane of the vide
screen. Peg positions are restricted to their respective planes.

Figure 4: *Peg Alignment Task Design*: Peg Position Resrticted to Plane

Experimental Protocol

TEST I:(Peg alignment without time constraint)

Each trial of TEST I was run as follows: The computer placed the target peg somewhere on the TARGET X-Z plane and projected the stereoscopic image using a particular interocular distance in the projection model. The subject would then be instructed to use the mouse to position the control peg so it was aligned directly below the target peg. Since the control peg is constrained to move only within the CONTROL X-Z plane, vertical alignment is guaranteed and not a factor in this study. The subject was allowed as much time as needed to get the two pegs lined up along the X and Z axes. When satisfied with the alignment, the subject would press a button on the mouse and the trial would be complete. For each trial the computer would record the X and Z target peg positions, the X and Z control peg positions, the time taken for the trial, and the interocular distance used for the trial.

For each of 9 subjects tested, 90 trials were run. Each trial tested a particular target peg position and projected the image with a particular interocular distance. All subjects were tested on the same distribution of target location/interocular distance pairs. Interocular distances ranging from 0 cm to 8 cm were tested, yielding a full range of stereopsis from pure monocular to enhanced stereo. Target locations were randomly mixed as were interocular distance trials. Thus the subjects had no way to predict the peg location in subsequent trials and had no knowledge of the interocular distance used for each projection. In fact, subjects were not informed that interocular distances were being varied during the experiment to ensure that such knowledge would not influence their performance.

TEST II: (Peg alignment with time constraint)

Each trial of TEST II was run identically to trials of TEST I in all ways except for the mode of trial termination. Rather than waiting for the pressing of a button to signal the end of the trial, the trial ended abruptly after 2.5 seconds had elapsed. Whereas in TEST I, subjects were given as much time as needed to align the pegs, in TEST II subjects were required to align the pegs as best as they could in the short interval provided. When the 2.5 second interval had elapsed, the target would disappear and data would be recorded for the trial. The subject would then be presented with a new target and be given a fresh 2.5 second interval. For each trial the computer would record the X and Z target peg positions, the X and Z control peg positions, and the interocular distance used for the trial.

For each of 8 subjects tested on TEST II, 90 trials were run. Each trial tested a particular target peg position and projected the image with a particular interocular distance. As in TEST I, all subjects were given identical distributions of target location/interocular distance pairs which included interocular distances ranging from 0 cm to 8 cm. Target locations were randomly mixed across trials as was the interocular distance used. Thus, a subject could not predict the peg location in subsequent trials and had no knowledge of the interocular distance used for each projection.

RESULTS

For each trial of each test, the following information was recorded: target peg positions in the horizontal (X), target peg positions in depth (Z), control peg positions in the horizontal (X), control peg positions in depth (Z), interocular distance used in the trial, and time elapsed during the trial.

Data Analysis

To get a meaningful indication of how user performance varied with interocular distance, the following statistical techniques were used. First, alignment errors for each of the X and Z axes were computed. These errors were calculated for each trial by subtracting the coordinates of the target peg from the coordinates of the control peg. The values were then grouped by the interocular distance so that performance could be correlated to eye separation used in the projection model.

Next, mean alignment errors and standard deviations of alignment errors were generated for each interocular distance. Mean alignment errors were first calculated across trials and then calculated across subjects. This analysis was performed separately for errors along the X and Z axes. These axes were kept uncoupled in the analysis because it was thought that interocular distance would affect performance in the depth axes differently than it would affect performance along the horizontal axes. Mean errors were graphed vs interocular distance for the result of TEST I as shown by Figures 5 through 8. Mean errors were graphed vs interocular distance for the result of TEST II as shown by Figures 9 and 10.

DISCUSSION

Looking first at the mean error analysis done on the data from TEST I, surprising relations between performance and interocular distance are revealed. Figure 5 shows a plot of mean alignment error (along the depth

axis) versus interocular distance across all subjects. As was expected, this plot shows a marked degradation in performance as interocular distance approaches zero. In fact, when 0 cm was used as the interocular distance in the projection model (corresponding to pure monocular vision), the mean error was roughly 10 times greater than the mean error seen when a physiologically typical interocular distance of 6 cm was used. These results strongly support the use of stereo projections over monocular projections to improve performance in depth perception tasks. It should be noted that the peg alignment task made use of three-dimensionally rendered and shaded pegs to assure the presence of rich monocular depth cues. During post-testing interviews, many subjects reported that size variation with depth was a primary depth cue used in alignment. Although subjects consciously used this monocular depth cue as a guide when performing this task, performance in trials with adequate stereopsis greatly surpassed performance in trials with little or no stereoscopic cues. This result suggests that stereoscopic vision enhances performance in depth perception tasks even when rich monocular depth cues are provided to the user.

If a curve is fit to the depth performance data displayed in Figure 5, a logarithmic relation between mean error and interocular distance emerges. Although this logarithmic relation predicts a dramatic increase in performance when interocular distance is less than 2 cm, very little change in performance is seen over most of the interocular distance range tested. In fact, there was no measurable increase in mean depth perception performance for interocular distances greater than 3 cm. Although the logarithmic curve was fit for the mean data across subjects, plotting each subject's performance data individually (as seen in Figure 6) shows that all subjects followed a similar pattern.

The lack of measurable performance change over most of the range of interocular distance tested has some interesting implications to the design of systems using stereoscopic projections. Results from TEST I suggest that any interocular distance greater than about 3 cm can be used in the projection model without compromising performance in depth perception tasks. This result alone is not of much significance unless there is some motivation for using particular values of interocular distance in the projection model. Such a

motivation does exist: the range of depths that can be presented to a user is greatly limited by a user's ability to fuse the image pair. If images are projected too far behind or in front of the plane of the screen, parallax values become so large that the user's visual system can no longer fuse the pair and a double image appears [13]. This is the same double image effect that occurs if you hold your finger too close to your eyes. Since the magnitude of parallax generated by the projection model is scaled by interocular distance, the smaller the value of interocular distance used, the greater the range of depth that can be achieved without loss of image fusion.

Another motivation for using the smallest possible value for interocular distance stems from the fact that although your brain perceives the object at some depth in front of or behind the screen, your eyes must remain *focused* on the plane of the screen to accurately see the images [12]. This contradiction between focal depth and perceived depth can cause user discomfort and fatigue. This effect can be reduced by using small values of parallax. Since the magnitude of parallax generated by the projection model is scaled by interocular distance, reducing the interocular distance in the projection model is an effective method of reducing this effect.

Turning attention next to mean alignment error along the horizontal axis, more surprising results are revealed. It was anticipated that little correlation between horizontal error and interocular distance would be seen because stereo depth perception is not required for horizontal alignment of the pegs. As shown by Figures 7 and 8, the results from TEST I suggest that this prediction is far from correct. In fact, a curve fit to the horizontal alignment data shows a logarithmic relation between performance and interocular distance similar to that for the mean alignment data. When an interocular distance of 0 cm was used in the projection model, the mean horizontal alignment error was about 10 times greater than the mean error seen when a physiologically typical interocular distance of 6 cm was used. Similar to the depth error data, the horizontal error data show no measurable increase in performance for interocular distances greater than 3 cm.

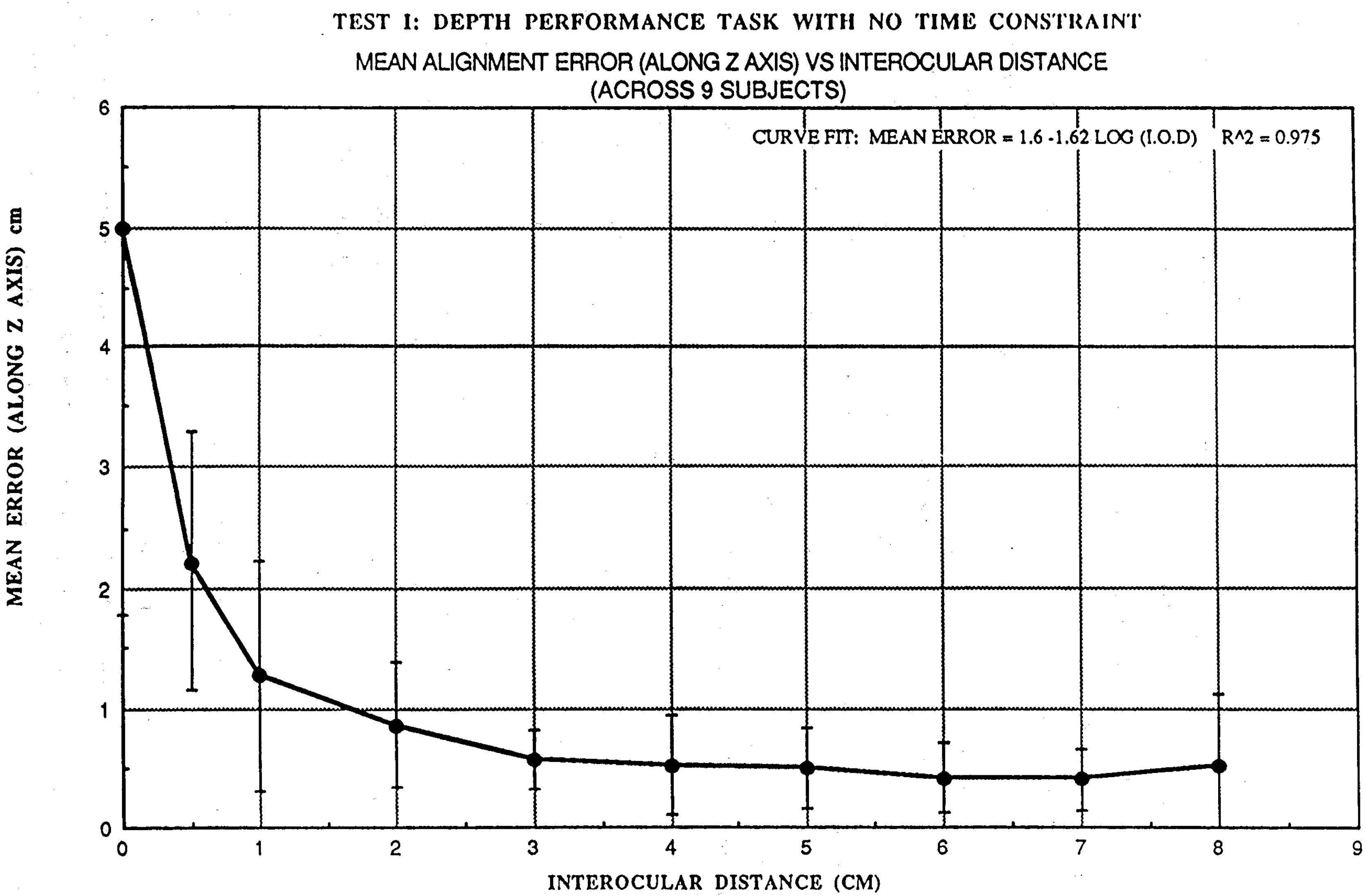


Figure 5: Mean Alignment Error (Along z Axis) vs Interocular Distance

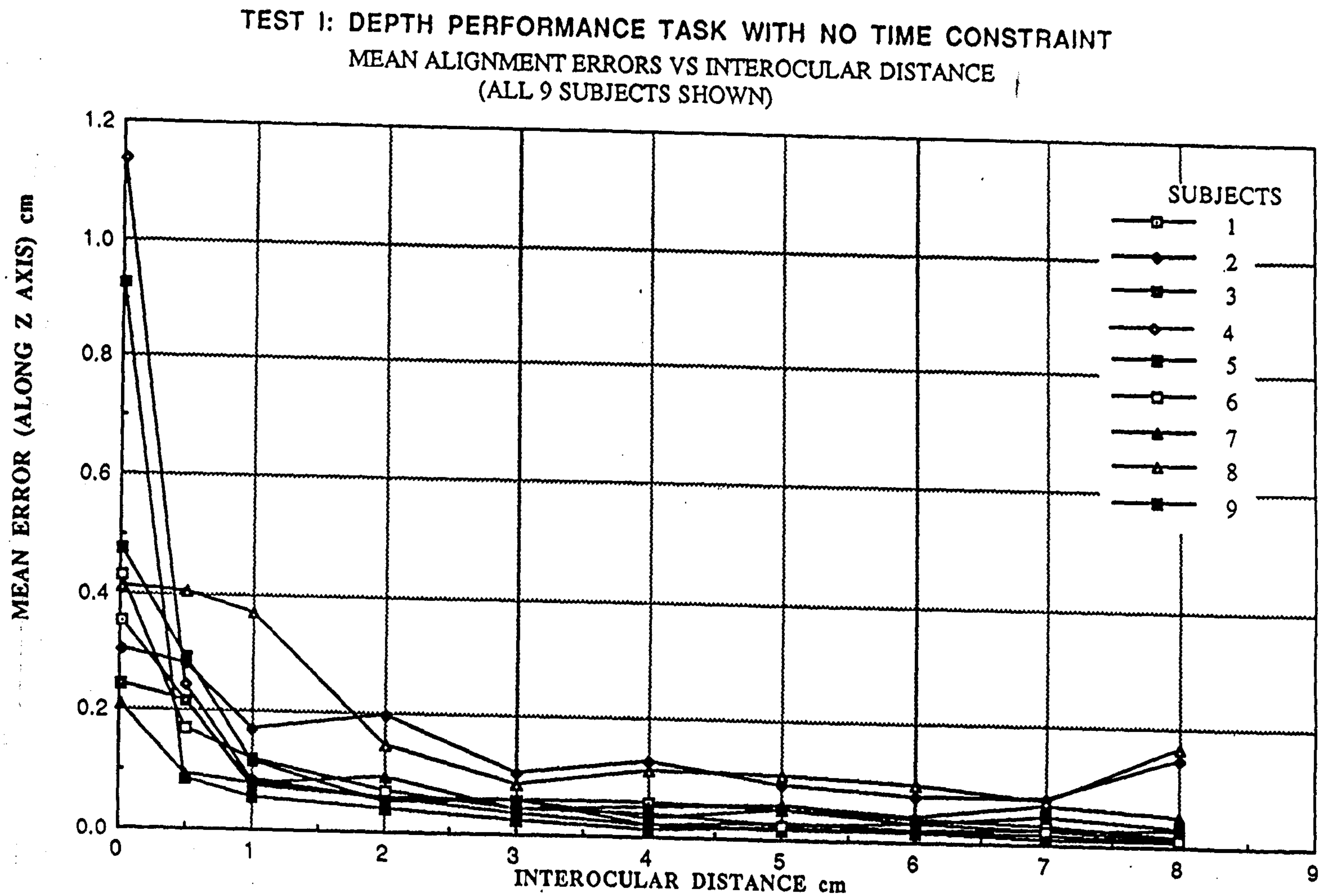


Figure 6: Alignment Error (Along z Axis) vs Interocular Distance For Each Subject

Why should interocular distance have an effect upon performance in horizontal alignment? The reason most likely results from the fact that we are projecting depth images normally perceived radially by the eyes onto a flat monitor. If we think of depth perception in radial coordinates rather than Cartesian coordinates and define a *viewing axis* as a radial *line of sight* from the center of the eyes to the object being viewed, we find a coupling between the horizontal axis and the depth along the line of sight axis. When viewing objects projected at the horizontal center of the screen, the viewing axis is aligned with the depth axis into the monitor. When viewing objects near the periphery, the viewing axis diverges from the depth axis. Thus errors in depth perception along the viewing axis will have a component in the horizontal Cartesian axis for targets that are not near the center of the screen. This hypothesis can be easily tested by comparing the results of those horizontal alignment trials with targets near the center of the screen to trials with targets near the periphery. If trials near the center show significantly higher performance than the trials near the periphery, it is likely that the projection of the radial image onto a flat screen is the source of horizontal errors.

To test this hypothesis, trials with a low interocular distance of 1 cm were examined to see if the poor stereopsis associated with this small interocular distance would result in greater horizontal errors near the periphery than near the center of the screen. Comparing trials across all subjects, the following results were found:

TABLE I. Alignment Error Correlated to Horizontal Location of Target on Screen

Horizontal Location of Target	Mean Horizontal Alignment Error
Trials within 1 cm of screen center	Mean Error = 0.026 cm
Trials within 1 cm of screen periphery	Mean Error = 0.14 cm

For trials with a small interocular distance of 1 cm, we see more than a 5-fold increase in alignment errors near the periphery of the screen compared to errors near the center of the screen. Thus, when stereopsis is poor,

horizontal alignment performance seems to be greatly influenced by target distance from the center of the screen. This result supports the hypothesis that horizontal alignment performance is influenced by stereopsis because images representing radial depth perception are projected onto a flat screen. Regardless of the cause of this effect, these results have interesting implications for the design of stereoscopic systems. It seems that a means of centering the image along the horizontal plane before performing visual tasks requiring horizontal alignment would enhance user performance.

Turning attention to the results from TEST II, we find very similar results to those revealed by TEST I. Whereas in TEST I subjects were given as much time as needed to align the pegs, TEST II allowed subjects only 2.5 seconds to complete the alignment. Not only did this speed constraint significantly increase the difficulty of the task, it prevented subjects from dwelling on the alignment task by giving them only enough time for very coarse positioning. Post-testing interviews confirmed that all subjects found TEST II to be significantly more challenging than TEST I and that most subjects felt that they were not given adequate time to complete the alignment task. Although TEST II posed an alignment task that was significantly more difficult than the task posed in TEST I, the relations between performance and interocular distance remained consistent with the results of TEST I. As shown in Figures 9 and 10, both the horizontal and depth analyses revealed characteristic logarithmic relations between performance and interocular distance. The consistency between results of TEST I and TEST II suggest that conclusions drawn from these simple depth tasks can be applied to tasks which span a wide range of paradigms and difficulties.

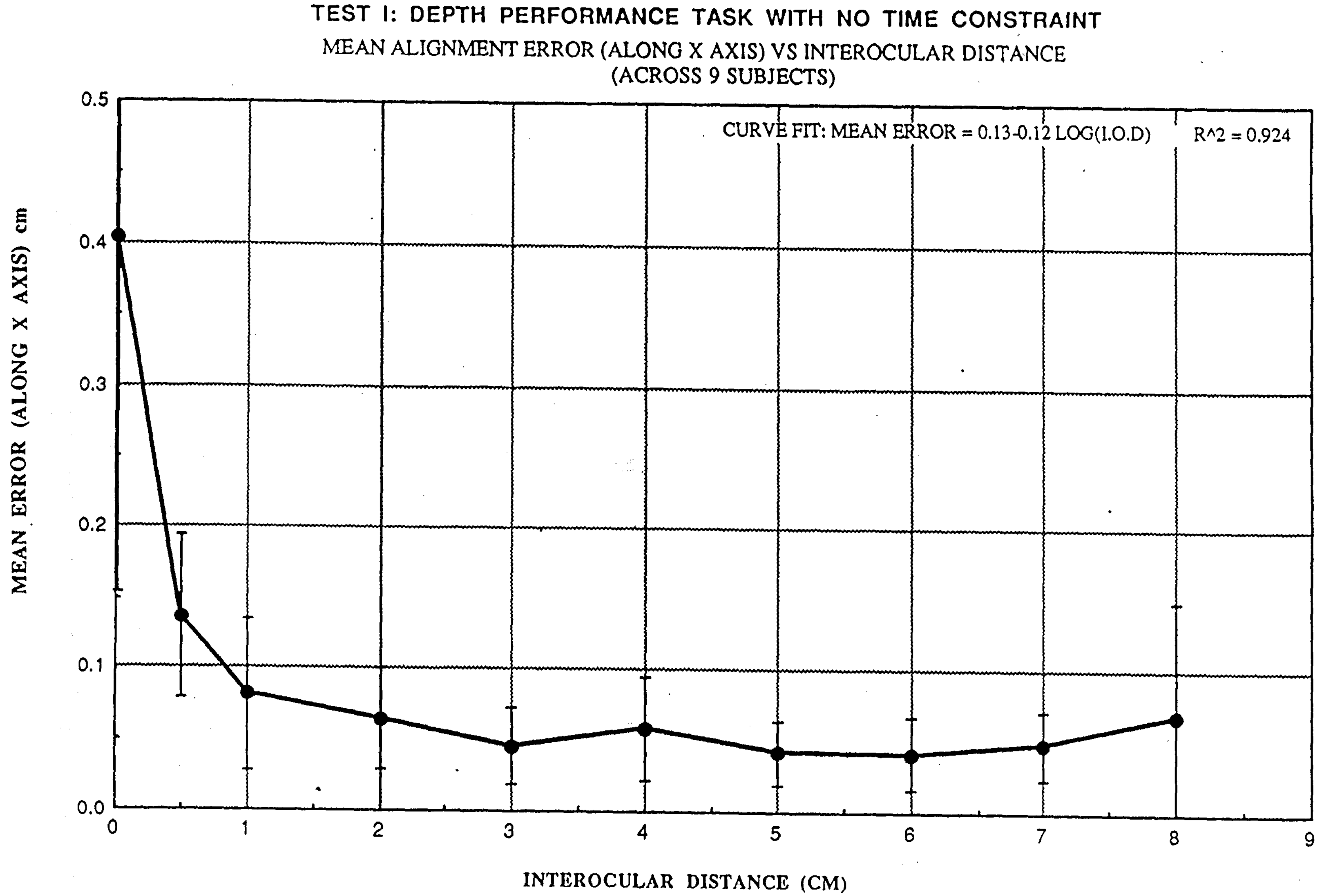


Figure 7: Mean Alignment Error (Along x Axis) vs Interocular Distance

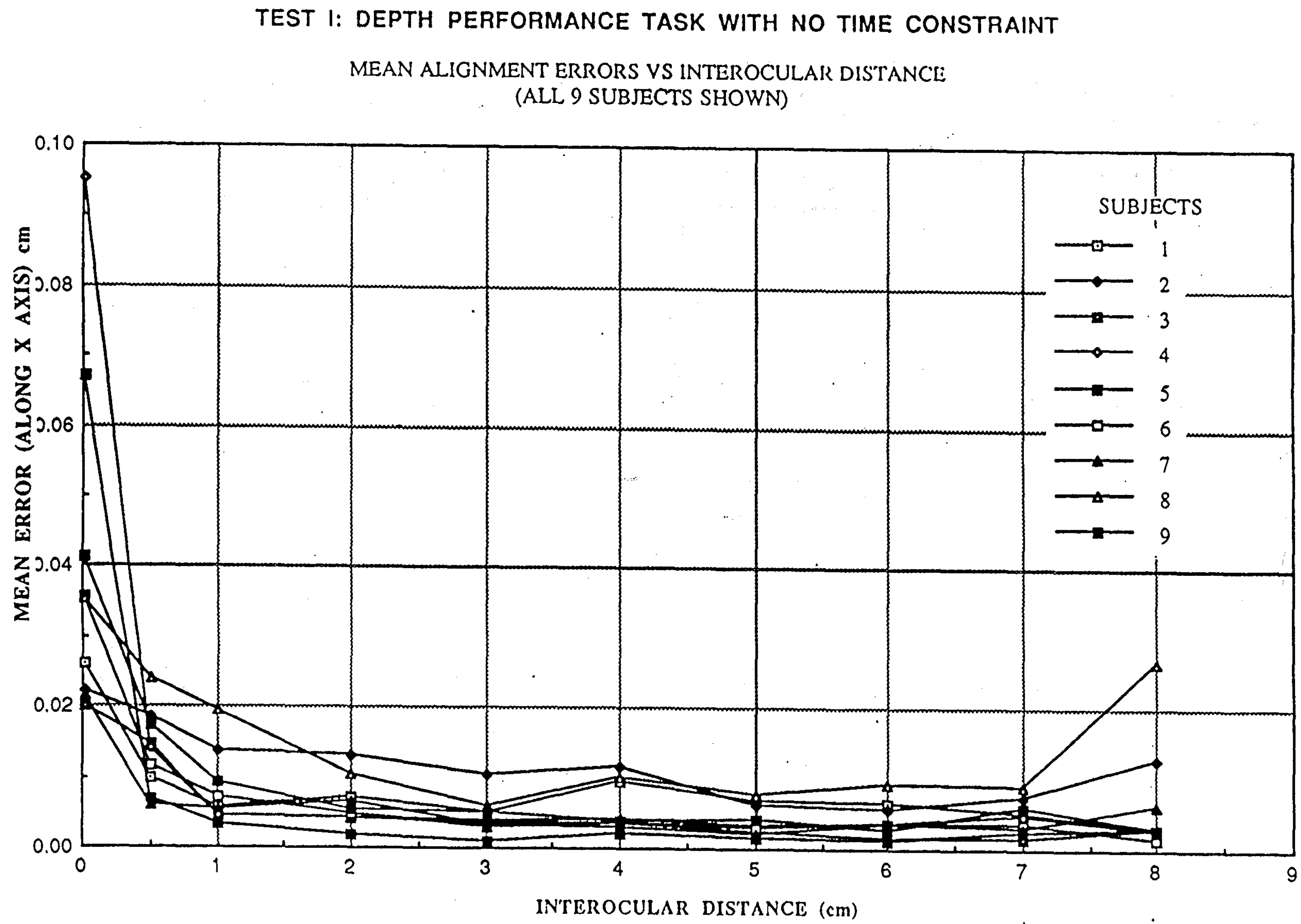


Figure 8: Alignment Error (Along x Axis) vs Interocular Distance For Each Subject

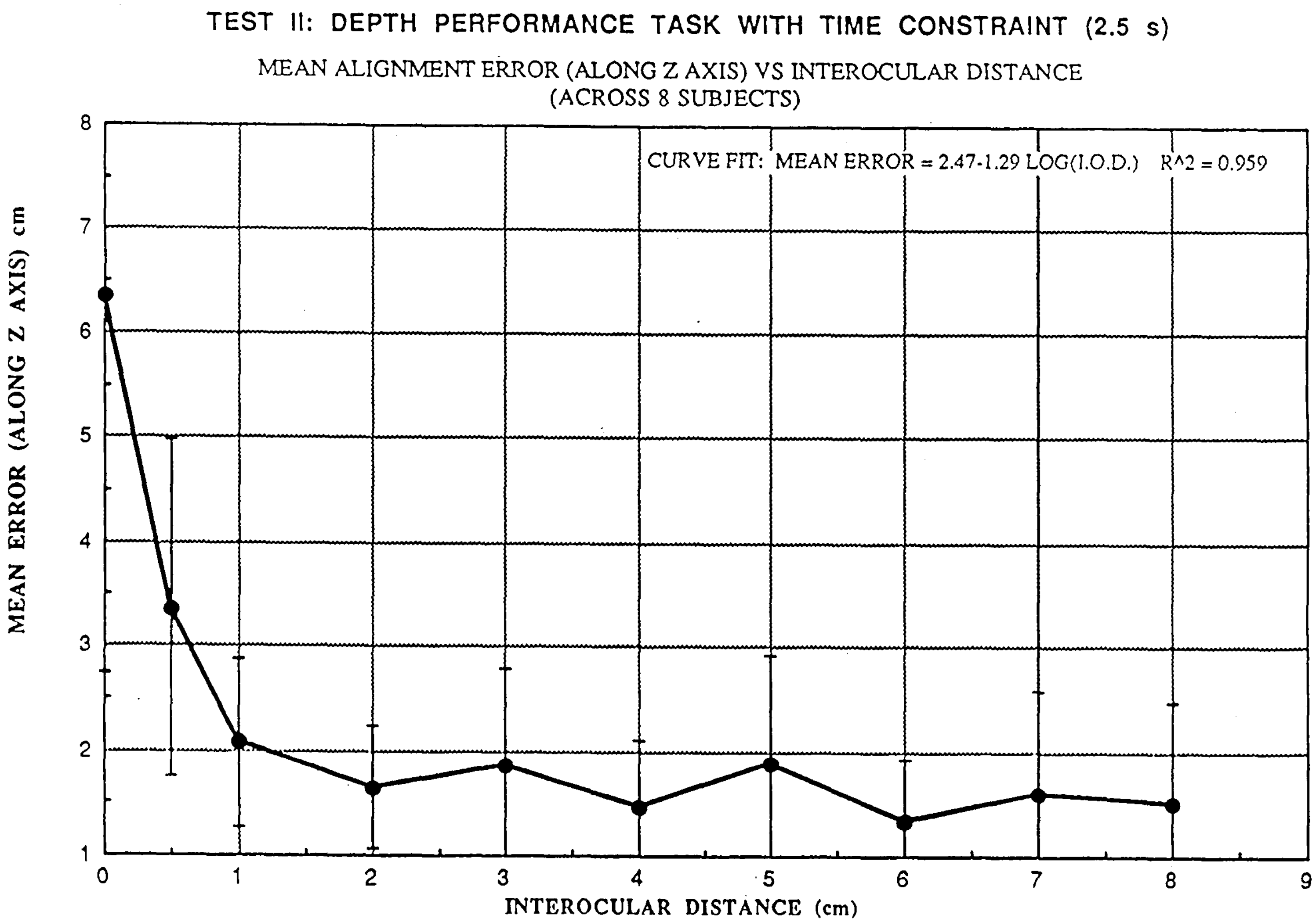


Figure 9: Mean Alignment Error (Along z Axis) vs Interocular Distance for Alignment Task With Time Constraint.

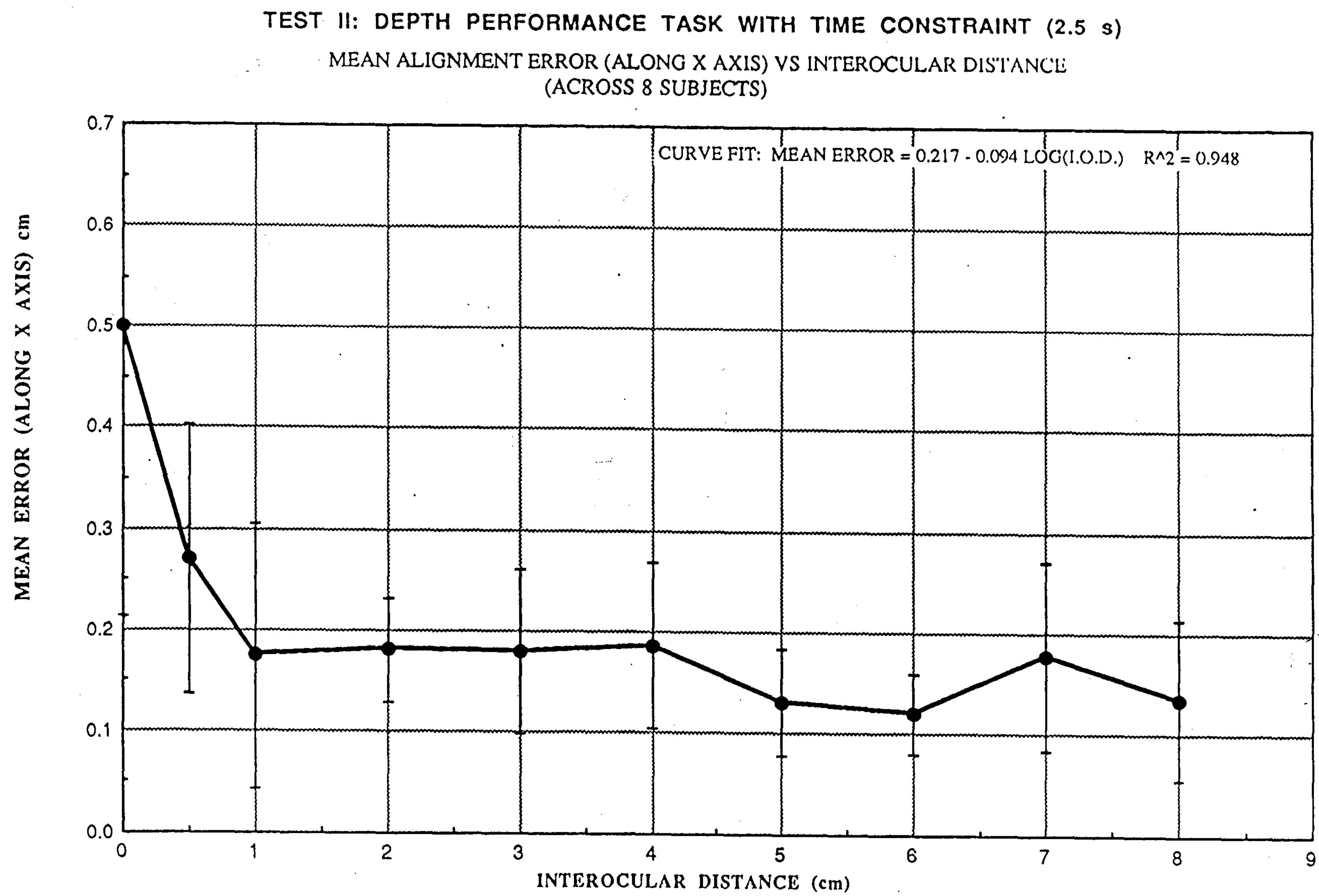


Figure 10: Mean Alignment Error (Along x Axis) vs Interocular Distance for Alignment Task With Time Constraint.

CONCLUSIONS

The following summarizes key points drawn from results of TEST I and TEST II.

1. When projected images are provided to a user performing a visual depth perception task in a telepresent or virtual environment, the use of stereoscopic projections results in a significant reduction of alignment errors over the use of pure monocular projections.

2. Although average physiological interocular distance is 6.3 cm, it was found that any distance of 3 cm or more was adequate to provide a user with maximal performance in the depth perception task. No statistically significant increase in performance could be correlated to increasing interocular distances greater than 3 cm. Since it is often beneficial to reduce the magnitude of parallax between the left and right images to increase the presentable depth range, reduce image fusion problems, and reduce operator fatigue, this result suggests that smaller than physiological interocular distances should be considered when implementing a stereoscopic vision system.

3. It was found that performance in horizontal alignment showed a very similar relation to interocular distance as performance in the depth axes. This result was surprising because stereopsis is not obviously required for horizontal alignment. Further investigation revealed that this effect was more prominent near the periphery of the screen than near the center. It is possible that this effect was the result of coupling between horizontal and depth axes due to the fact that line of sight depth perception, a radial phenomenon, was projected onto a flat monitor. This result suggests that some means of centering the target before performing horizontal alignment tasks would improve performance in both stereoscopic and monocular vision systems.

4. It was found that the self-paced depth perception task presented in TEST I yielded very similar results to the time pressured, more difficult depth perception task presented in TEST II. This result suggests that the conclusions drawn from these tests are largely independent of the difficulty of the task and may be applicable to a wide range of depth perception tasks.

REFERENCES

[1] Boff, K.R. and J.E. Lincoln, Engineering Data Compendium: Human Perception and Performance. AAMRL, Wright-Patterson AFB OH, 1988.

[2] Donelson, Sarah, 1988 Anthropometric Survey of U.S. Army Personnel, U.S. Army Technical Report Natick/TR-91/040, Natick MA, 1988, 198-199.

[3] Kama, W.N. and R.C. DuMars, Remote Viewing: A Comparison of Direct Viewing, 2-D and 3-D Television, AMRL-TDR-64-15, Wright-Patterson AFB OH, 1964.

[4] Rupert, P. and D. Wurzburg, "Illumination and television considerations in teleoperator systems," in Heer, E. (Ed.) Remotely Manned Systems, California Institute of Technology, CA, 1973, 219-228.

[5] Freedman, L.A., W.H. Crooks, and P.P. Coan, "TV requirements for manipulation in space," Mechanisms and Machine Theory, 12, 5, 1977, 425-438.

[6] Pepper, R.L. and P.K. Kaomea, "Teleoperation: Telepresence and Performance Assessment," Annual International Ergonomics Society Meeting, Teleoperation and Control, July 1988.

[7] Pepper, R.L., D.C. Smith, and R.E. Cole, "Stereo TV improves operator performance under degraded visibility conditions," Optics Engineering, 20, 4, 1981, 579-585.

[8] Pepper, R.L. R.E. Cole, E.H. Spain, and J.E. Sigurdson, "Research issues involved in applying stereoscopic television to remotely operated vehicles," Proceedings of SPIE, 402 - Three Dimensional Imaging, 1983, 170-173.

[9] Kim, W.S., S.R. Ellis, M. Tyler, and L. Stark, "Visual enhancements for telerobotics," IEEE Proceedings of the International Conference on Cybernetics and Society, 1985, 819-823.

[10] Ellis, S.R., W.S. Kim, M. Tyler, M.W. McGreevy, and L. Stark, "Visual Enhancement for Telerobotics: perspective parameters," IEEE Proceedings of the International Conference on Cybernetics and Society, 1985, 815-818.

[11] Kim, W.S., S.R. Ellis, B. Hannaford, M. Tyler, and L. Stark, "A quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks," IEEE Transportation Systems, Man, and Cybernetics, SMC-17, 1, 1987, 61-72.

[12] Lipton, Lenny, The Creation of Software for Stereoscopic Computer Graphics, StereoGraphics Corporation, San Rafael CA, 1988.

[13] Akka, Robert, Software Concepts for Stereoscopic 3D Graphics, StereoGraphics Corporation, San Rafael CA, 1991.

[14] Akka, Robert, Automatic Software Control of Display Parameters for Stereoscopic Graphics Images, StereoGraphics Corporation, San Rafael CA, 1991.

U. S. Government Printing Office 1994 550-057/00007